

## PRISTOP K IZGRADNJI DIGITALNIH ARHIVOV

*Alenka Kavčič-Čolić\**

UDK: 004:930.25

*Alenka Kavčič-Čolić: Pristop k izgradnji digitalnih arhivov. Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Zbornik referatov z dopolnilnega izobraževanja, Maribor 4/2005, št. 1, str. 480-490.*

*Izvirnik v slovenščini, izvleček v slovenščini in angleščini, povzetek v angleščini.*

Izhajajoč iz mednarodnih analiz in priporočil skupine RLG/OCLC ter UNESCO-vih Smernic za trajno ohranjanje digitalne dediščine prispevek obravnava značilnosti zaupanja vrednega digitalnega arhiva in predstavlja pomembnejše vidike, ki bi jih morali upoštevati pri njihovi gradnji, kot so: vprašanje odgovornosti, kriteriji selekcije, sodelovanje z avtorji oz. založniki digitalnih virov, upravljanje z avtorskimi pravicami, odločanje o strategiji trajne zaščite digitalnih virov, izbira struktur metapodatkov in drugo.

UDC: 004:930.25

*Alenka Kavčič-Čolić: An Approach to the Development of Digital Archives. Technical and Field Related Problems of Traditional and Electronic Archiving. Conference Proceedings, Maribor 4/2005, No. 1, pp. 480-490.*

*Original in Slovenian, abstract in Slovenian and English, summary in English.*

Based on international analyses and recommendations by RLG/OCLC and UNESCO Guidelines for Preservation of Digital Heritage an overview of the attributes of trusted digital archives is given and some aspects that should be taken into consideration in their development are presented. These aspects comprise: responsibility issues, selection criteria, cooperation with providers, copyright management, selection of long-term preservation strategies, metadata, etc.

### 1. UVOD

Splošno je znano, da so knjižnice in arhivi zadolženi za ohranjanje narodne kulturne dediščine. To odgovornost so prevzeli pred mnogimi leti in na njej so gradili svojo podobo zanesljivih institucij, katerih glavna skrb je, da bodo avtorska dela in dokumenti za vedno shranjeni oz. dostopni javnosti. V ta namen so razvili več sistemov shranjevanja, organizacije in lokalizacije gradiva. Vendar njihova dolgoletna tradicija temelji le na ohranjanju analognih del ne pa tudi digitalnih. Digitalne<sup>1</sup> publikacije in dokumenti imajo značilnosti, ki zahtevajo določene pogoje za njihovo ohranjanje. Zato se lahko vprašamo, ali so ravno te ustanove, katerim bi zaupali dragocene dokumente in dela, ki so bila ustvarjena na papirju ali drugi analogni obliki, sposobne trajno<sup>2</sup> ohranjati tudi digitalne zapise. Kakšne pogoje mora

---

\* *Alenka Kavčič-Čolić, vodja Bibliotekarskega raziskovalnega centra Narodne in univerzitetne knjižnice, Turjaška 1, Ljubljana, Slovenija.*

<sup>1</sup> *Termin "digitalne publikacije" je v prispevku uporabljen kot protipomenka "analognih" in sopomenka "elektronskih" publikacij. Digitalne publikacije zajemajo vse elektronsko gradivo, ustvarjeno v elektronski obliki ali ustvarjeno na podlagi analognega originala. V tem smislu so uporabljeni termini: digitalno gradivo, digitalna zbirka, digitalni vir in digitalni dokument.*

<sup>2</sup> *Dobesedni prevod izraza "long-term preservation" je "dolgodobno oz. dolgoročno ohranjanje". Kljub temu smo se odločili za uporabo izraza "trajno ohranjanje", ker izhajamo iz dejstva, da želimo vire trajno ohraniti, čeprav je pojem trajnosti zelo relativen.*

zadovoljiti ustanova, da bi bila vredna našega zaupanja tudi pri shranjevanju digitalnega gradiva? Obljube in zagotovila ne zadostujejo.

Že leta 1994 je Komisija za trajno ohranjanje in dostop (*Commission on Preservation & Access*) iz Washingtona, v sodelovanju z Skupino raziskovalnih knjižnic za arhiviranje digitalnih informacij (*RLG Taks Force on Archiving of Digital Information*) ugotovila potrebo po definiciji pogojev, ki jih morajo izpolnjevati zanesljivi digitalni repozitoriji<sup>3</sup>. V njihovem poročilu iz leta 1996 (Garret & Waters, 1996, str.5) je med ostalim naštetu, da je za trajno ohranjanje digitalnega gradiva nujno potrebna ustrezna infrastruktura za podporo distribuiranega sistema digitalnih repozitorijev; zadostno število zaupanja vrednih organizacij, ki bi bile sposobne shranjevati, migrirati in omogočiti dostop do digitalnih zbirk; ter uvajanje prakse dodeljevanja certifikatov, ki bo povečal zanesljivost digitalnih arhivov.

Leta 2000 je RLG (*Research Libraries Group*) začela sodelovati z OCLC (*Online Computer Library Center*) in rezultat tega sodelovanja je poročilo: *Trusted digital repositories: attributes and responsibilities*, ki ga je RLG objavila leta 2002. Publikacija je zelo pomembna, ker obravnava značilnosti in vse odgovornosti, ki jih morajo prevzeti zaupanja vredni digitalni arhivi (*trusted digital repositories*). Med avtorji so znani strokovnjaki s področja trajnega ohranjanja digitalnega gradiva. Catherine Lupovici, vodja oddelka za digitalno knjižnico v francoski nacionalni knjižnici, je sodelovala v evropskem projektu NEDLIB (*Networked European Deposit Library*)<sup>4</sup>, ki je imel za cilj razvoj depozitnega sistema za trajno ohranjanje digitalnih publikacij. V okviru projekta je prispevala k oblikovanju modela metapodatkov za trajno ohranjanje v digitalnem repozitoriju<sup>5</sup>. Deborah Woodyard, tedanja koordinatorka za trajno ohranjanje digitalnega gradiva v *British Library*, je zelo aktivna v RLG tudi na področju definicije metapodatkov za trajno ohranjanje. Colin Webb, direktor službe za trajno ohranjanje pri avstralski nacionalni knjižnici, je avtor poznejšega dela iz leta 2003, *Guidelines for the preservation of digital heritage* (v nadaljnjem besedilu Smernice), ki je nastalo po naročilu UNESCO. Omenjene Smernice po vsebini sledijo UNESCO-vi Listini o trajnem ohranjanju digitalne dediščine, ki je bila sprejeta na 32. seji Generalne Konference UNESCO oktobra 2003, in obravnavajo vse vidike pri gradnji digitalnih arhivov.

V nadaljevanju so našteje značilnosti in odgovornosti zaupanja vrednih digitalnih arhivov, predstavljeni so tudi pomembnejši vidiki, ki bi se morali na osnovi navedenih priporočil in smernic upoštevati pri njihovi gradnji.

## 2. ZNAČILNOSTI IN ODGOVORNOSTI ZAUPANJA VREDNIH DIGITALNIH ARHIVOV

Na osnovi poročila RLG/OCLC iz leta 2002 je glavna naloga zaupanja vrednega digitalnega arhiva (*trusted digital repository*)<sup>6</sup> omogočiti določeni skupnosti v sedanjosti in prihodnosti zanesljiv in trajen dostop do ohranjenih digitalnih virov. V

---

<sup>3</sup> V besedilu je pomen arhiva širši od repozitorija. Medtem ko digitalni repozitorij označuje računalniško aplikacijo za shranjevanje digitalnega gradiva, digitalni arhiv predstavlja organizacijo in sistem za trajno ohranjanje, vključno z dejavnostmi, ki so potrebne, da bi se to ohranjanje izvajalo. Digitalni arhiv je lahko del knjižnice, arhiva ali muzeja.

<sup>4</sup> Domača stran projekta <http://www.kb.nl/coop/nedlib/> (pregledano 21. 9. 2004.)

<sup>5</sup> Lupovici, Catherine and Masanès, Julien: *Metadata for Long Term Preservation*. - Den Haag: Koninklijke Bibliotheek, 2000. - (NEDLIB Report series; 2).

<sup>6</sup> V literaturi se pogosto mešata pojma arhiv in repozitorij. Pri nazivu *Trusted digital repository* je uporabljen termin repozitorij za označevanje arhivskega sistema.

istem poročilu se navajajo njegove glavne značilnosti (Research Libraries Group, 2002, str. 13):

- 2.1 **Njegov repozitorij je usklajen z referenčnim modelom OAIS (*Open Archival Information System*).**
- 2.2 **Odgovoren je za svoje poslovanje.** Njegovo poslovanje mora biti v skladu dobre prakse in zagotavljati mora zanesljivost in trajnost. Slediti mora domačim in/ali mednarodnim standardom glede infrastrukture, potrebne za delovanje repozitorija, varnostnega preslikovanja in drugih varnostnih ukrepov. Mora imeti jasno definirano politiko in strategijo glede nabave, vzdrževanja, dostopa do gradiva in sodelovanja z vlagatelji vsebin. Poslovanje mora biti transparentno javnosti.
- 2.3 **Organizacijsko je zanesljiv.** Organizacija, ki prevzame vlogo digitalnega arhiva, mora pokazati, da je zanesljiva. Cilj trajnega ohranjanja digitalnega gradiva mora biti jasno naveden v njeni strategiji. Dejavnosti, navedene v njenem ustanovitvenem aktu, so usklajene z odgovornostjo, ki jo prevzema na področju ohranjanja digitalnih virov. Poleg tega mora imeti ustrezno infrastrukturo, strokovne kadre za gradnjo in vzdrževanje digitalnega arhiva. Mora tudi imeti formalni sukcesijski načrt.
- 2.4 **Finančno je zanesljiv.** Organizacija, ki opravlja vlogo digitalnega arhiva, mora imeti zanesljiv dolgoročni in kratkoročni poslovni načrt, v katerem je predviden redni pritek finančnih sredstev za dejavnost, razvoj in vzdrževanje digitalnega repozitorija.
- 2.5 **Zagotavlja tehnološko in procedurno ustreznost pri trajnem ohranjanju.** Digitalni arhiv mora predvideti različne strategije za trajno ohranjanje in v ta namen mora zagotoviti ustrezno informacijsko-tehnološko infrastrukturo ter natančen načrt postopkov in scenarijev pri izvajanju omenjenih strategij.
- 2.6 **Zagotavlja visoko sistemsko varnost svojega repozitorija.** Digitalni arhiv mora poskrbeti za politiko in prakso pri zagotavljanju varnosti digitalnega gradiva. Rešiti mora problem preslikovanja, avtoriziranja in podvajanja podatkov, zagotoviti mora sistem za overovitev dostopa do repozitorija, za ustvarjanje varnostnih kopij gradiva ter zelo zanesljiv požarni zid. Mora imeti načrt za primer izpada sistema, izgube ali spremembe podatkov ter možnost za povrnitev izgubljenih podatkov.
- 2.7 **Odgovoren je za vse vidike svojega delovanja.** Zaupanja vredni arhiv mora prevzeti odgovornost za svoje delovanje, za izbiro gradiva, za upravljanje s celotnim življenjskim ciklusom digitalnega gradiva, za izpolnjevanje pričakovanj interesnih skupnosti, za urejanje avtorsko-pravnih vidikov in za vse stroške pri trajnem ohranjanju.

Vse našete značilnosti so pomembne in se upoštevajo pri dodeljevanju certifikatov. Skupina RLG/OCLC se zavzema za uvajanje sistema certificiranja, ki bi potekal postopoma v več fazah in bi se po nekaj letih prenavljalo (Research Libraries Group, 2002, str. 33). V Nemčiji je *Deutsche Initiative für Netzwerkinformation* leta 2003 začel z dodeljevanjem DINI-jevega certifikata (DINI-Certificate, 2003) organizacijam, ki zadovoljujejo osnovne kriterije za trajno ohranjanje digitalnega

gradiva. Podlaga za analizo so vprašalniki, ki jih zainteresirane arhivske organizacije izpolnijo in posredujejo DINI ter plačilo določenega zneska za kritje stroškov<sup>7</sup>.

### 3. PRISTOP K IZGRADNJI DIGITALNIH ARHIVOV

V UNESCO-vih smernicah se upoštevajo vse našete značilnosti in odgovornosti digitalnih repozitorij. Ker pa so po vsebini priporočila za izgradnjo digitalnih arhivov, so opisani postopki in vidiki, ki jih moramo upoštevati pri njihovi izgradnji. Te postopke lahko povzamemo v naslednjih točkah:

- prevzemanje odgovornosti za trajno ohranjanje digitalnega gradiva,
- kriteriji izbire digitalnih virov,
- sodelovanje z avtorji/založniki digitalnih virov,
- urejanje avtorsko-pravnih vidikov,
- izbira metodologije ohranjanja digitalnih virov,
- usklajenost arhivskega sistema z referenčnim modelom OAIS,
- izbira primernih metapodatkov.

#### 3.1 PREVZEMANJE ODGOVORNOSTI ZA TRAJNO OHRANJANJE DIGITALNEGA GRADIVA

O odgovornostih je bilo govora v prejšnjem poglavju. Zelo pomembno je, da se organizacija, ki se odloča za trajno ohranjanje, zaveda vseh odgovornosti, ki jih pri tem prevzema. Vlogo digitalnega arhiva lahko prevzame ustanova, ki že ima izkušnje pri trajnem ohranjanju kulturne dediščine, npr. knjižnica, arhiv ali muzej. Lahko pa je tudi institucija, ki na komercialni osnovi ponuja storitev trajnega ohranjanja določene digitalne zbirke. Tega v Sloveniji še nimamo. Seveda so odgovornosti različne glede na vrsto gradiva, čas in način ohranjanja. Pomembno pa je, da nekdo prevzame odgovornost.

Ta odgovornost je lahko tudi razdeljena med različne institucije. Tako na primer nacionalne knjižnice skrbijo za pisno kulturno dediščino v digitalni obliki, arhivi za arhivsko digitalno gradivo, univerze za elektronske disertacije ali magistrske naloge, muzeji pa za digitalne publikacije ali digitalizirane podobe svojih eksponatov.

Odgovornost je lahko tudi časovno omejena. Elektronske založbe lahko prevzamejo odgovornost v času trajanja njihovega elektronskega založniškega programa. S prenehanjem delovanja ali z odstranjevanjem starih arhivov lahko to odgovornost prenesejo na drugo institucijo.

Priporočljivo je, da ena od organizacij prevzame vodilno vlogo oz. da koordinira dejavnosti na svojem področju ohranjanja. Optimalno bi bilo, da se različne ustanove povežejo v mrežo in redno usklajujejo svoje izkušnje in odgovornosti.

V smernicah se priporoča, da jasno definiramo gradivo, za katero prevzemamo odgovornost.

---

<sup>7</sup> Cena pridobivanja DINI-jevega certifikata za člane iz neprofitnih organizacij je 50€, iz profitnih organizacij pa 150€. Cena za nečlane iz neprofitnih organizacij je 100 €, iz profitnih pa 250€.

### 3.2 KRITERIJI IZIBIRE DIGITALNIH VIROV

Izbira digitalnih virov je zelo težka naloga, predvsem zaradi vse večjega števila digitalnih publikacij, samostojnih elektronskih založnikov, specifičnih formatov in podobno. Danes težko predvidimo ali bo ohranjanje določenega digitalnega vira pomembno za prihodnost. Težko tudi ugotavljamo, ali je digitalno delo, ki ga želimo ohraniti, avtentično. Ohranjanje ne sme potekati mimo dovoljenj imetnikov avtorskih pravic.

Pomembno je, da pri izbiri digitalnih virov dajemo prednost publikacijam, ki so nastale v digitalni obliki ter vsebinam, ki so znanstvenega, strokovnega ali drugega nacionalnega pomena. Seveda je odločitev tudi odvisna od razpoložljivih kadrov in zmožnosti sistema za shranjevanje.

V smernicah se priporoča, da se pred začetkom zbiranja digitalnega gradiva formalno določijo pravila pri njihovi izbiri. Omenjena pravila bi morala pomagati rešiti naslednje dileme:

- kako ravnati v primeru, ko je isti vir na več nosilcih (npr. tiskano na papirju, na magnetnih trakovih, na CD-ROM-u in/ali na medmrežju);
- kako ravnati v primeru, ko obstaja novejša različica (izdaja) digitalnega dokumenta, ki vsebuje manjše popravke in te ne vplivajo na vsebino;
- kako ravnati v primeru, ko imamo isti digitalni dokument v različnih formatih (npr. HTML, MS Word, PDF);
- po kakšnih postopkih lahko ugotovimo verodostojnost originalnih digitalnih dokumentov.

Pri izbiri nam pomaga, če lahko predvidimo, kdo bo uporabnik digitalnega dokumenta v prihodnosti in ali lahko zagotovimo vsa avtorsko-pravna dovoljenja za njegovo ohranjanje. Selekcija naj bi temeljila na vrednosti gradiva, kar pomeni, da stroški njegovega prevzemanja, ohranjanja in vzdrževanja ne bi smeli biti večji od njegove vrednosti.

### 3.3 SODELOVANJE Z AVTORJI/ZALOŽNIKI DIGITALNEGA GRADIVA

V smernicah je velik poudarek na sodelovanju z avtorji in založniki digitalnega gradiva. Ti nam pomagajo, da izvemo podatke o vseh udeležencih v procesu publikacije in pridobimo tudi vse informacije, ki so pomembne za trajno ohranjanje gradiva. Če lahko sledimo publikaciji od njenega nastanka, lahko tudi vplivamo na formate za skladiščenje in na standardizacijo postopkov pri objavi publikacije. Smernice tudi priporočajo, da se digitalni viri ohranjajo v formatih odprtega (nekomercialnega) tipa, ki lahko olajšajo procese njihove konverzije v novejšo ali druge različice. Obenem imamo zagotovilo o avtentičnosti in celovitosti pridobljenega gradiva.

Poleg naštetega nam sodelovanje z avtorji in založniki pomaga pri razumevanju pomembnosti in načina delovanja njihovih publikacij, značilnosti ciljne publike in načrtovanega trenda razvoja informacijske tehnologije založnikov ter želje samih avtorjev.

### 3.4 UREJEVANJE AVTORSKO-PRAVNIH VIDIKOV

Ob arhiviranju digitalnih virov moramo zagotoviti tri zvrsti avtorskih pravic:

- pravica do prevzemanja (preslikovanje digitalnega vira v arhiv),
- pravica do dostopa (preslikovanje digitalnega vira na delovni prostor računalnika v mreži),
- pravica do uporabe strategij za dolgoročno ohranjanje (pravica do podvajanja kopije digitalnega vira, do posegov v dokument zaradi uporabe strategij za dolgoročno ohranjanje - konverzije, migracije ...).

Nacionalkam in arhivskim institucijam lahko zakonodaja o obveznem izvodu publikacij omogoča prvo pravico. Druge dve pa morajo biti posebej zagotovljene s pogodbo ali licenco.

Vsekakor pa je pomembno, da so digitalni viri ohranjeni brez zaviralcev ali omejevalcev dostopa, saj jih je lažje uporabiti v prihodnosti.

Leta 1999 sta CENL in FEP pripravila dokument "*Code of practice for the voluntary deposit of electronic publications*" (CENL & FEP, 1999), kjer sta določila tudi osnove za sodelovanje z založniki in avtorji pri depozitu digitalnih virov.

### 3.5 IZBIRA METODOLOGIJE OHRANJANJA DIGITALNIH VIROV

Pri dobrem programu ohranjanja je zelo pomembno, da pred vnosom digitalnega vira v arhiv določimo možne strategije za njegovo trajno ohranjanje. Izbira določene strategije ohranjanja je zelo odvisna od dovoljenj imetnikov avtorskih pravic in od značilnosti samega digitalnega vira. Najprej moramo ugotoviti, kaj želimo ohraniti pri digitalnem viru.

Digitalni vir sestavlja več nivojev:

- **Fizični nivo** predstavlja digitalni zapis oz. sama oznaka na nekem fizičnem nosilcu, ki je lahko CD-ROM, disketa, magnetni trak ali kaj drugega;
- **Logični nivo** je sintaksa ali programski jezik, v katerem je ohranjen zapis. Pri novejših različicah računalniških programov se ta lahko tudi povsem spremeni. Enako pri konverzijah v druge računalniške aplikacije;
- **Konceptualni nivo** nam omogoča dostop do vsebine digitalnega vira. Na primer: ne glede na to, ali je datoteka v PDF ali v Word-u, jo mi lahko preberemo;
- **Nivo posebnih značilnosti digitalnega vira** - to so vse tiste značilnosti, ki naredijo digitalni vir tak, kot je, in ki so pomembne za uporabnika. Na primer: ni vseeno, če je pri digitalizirani sliki muzejskega eksponata ali umetnine zmanjšana resolucija ali razpon barv oz. če besedilo, ki je v PDF formatu ne moremo nikakor spreminjati ali če je ohranjeno kot digitalna slika in ne moremo iskati po njem.

Pri izbiri strategiji za trajno ohranjanje upoštevamo, kaj je bolj pomembno ohraniti: podatkovni tok, ki sestoji iz fizičnega in logičnega nivoja, ali dostop do vsebine digitalnega vira, katerega sestavljata konceptualni nivo oz. nivo posebnih značilnosti.

### 3.5.1 OHRANJANJE PODATKOVNEGA TOKA

Pri skladiščenju podatkovnega toka mora repozitorij v digitalnem arhivskem sistemu imeti zadostno kapaciteto, da prenese tudi večje količine podatkov. (V Tabeli 1 je nekaj primerov digitalnih virov in njihove potrebe po računalniškem spominu.) Omogočiti mora varnostno podvajanje podatkov, v primeru nesreč pa mora imeti tehnično podporo za hitro reagiranje. Poleg tega mora repozitorij biti sposoben upravljati z redundantnim digitalnim gradivom in avtomatsko preverjati napake.

10 vrstic besedila v ASCII	1 Kb
Statična spletna stran v HTML	20 Kb
Povprečna spletna stran na internetu (50% slik in 50% besedil)	11 Kb
Digitalni članek v PDF	50-500 Kb
200-2.000 revij s 340 člankov po 10 str. v času 10-ih let v PDF	1 Tb
2.000-20.000 knjig po 100 str. v PDF	1 Tb
Digitalizirana publikacija v TIFF	20-200 Mb
5-50 str. v TIFF	1 Gb
50-500 knjig po 100 str. v TIFF	1 Tb
Industrijska baza podatkov	1 Pb
Digitalizirana publikacija v JPEG	50 Kb-1 Mb
CD-ROM	Max. 650-700 Mb

**Tabela 1: Velikost različnih digitalnih virov**

Zaradi omejene življenjske dobe fizični nosilci nikakor ne morejo biti predmet trajnega ohranjanja. Vemo, da imajo magnetni nosilci (trdi disk, trakovi) kljub veliki kapaciteti (do 200 Gb ali več) življenjsko dobo ca. 5 let, optični (CD-ROM/DVD) pa celo do 100 let<sup>8</sup>, vendar so njihove kapacitete precej manjše (ca. 4 Gb). V vsakem primeru je optimalno načrtovati prenos na nove nosilce vsakih pet let.

### 3.5.2 OHRANJANJE DOSTOPA DO DIGITALNEGA VIRA

Pri ohranjanju dostopa ni enotnih rešitev. V vsakem primeru obstaja nevarnost izgube podatkov. Poleg tega ta pristop lahko zahteva visoke finančne stroške.

V Smernicah se obravnavajo mogoče strategije za trajno ohranjanje dostopa do digitalnega vira. Te so:

- **ohranjanje tehnologije** (muzejski pristop);
- **reinkarnacija starih strojev** (s pomočjo dodatnih čipov na hardveru);

<sup>8</sup> V okviru češke nacionalne knjižnice Klementine deluje podjetje, ki opravlja zunanje izvajanje digitalizacije za knjižnico. V tem podjetju imajo tehnologijo, s pomočjo katere letno merijo kvaliteto CD-ROM-ov, na katerih shranjujejo digitalizirano gradivo. Meritve so pokazale, da se njihova kvaliteta s časom manjša do nekaj odstotkov na leto.

- **migracija** (prenos podatkov iz ene generacije softvera/hardvera v drugo: menja se logični format, vsebina pa je lahko predstavljena z novo strojno in/ali programsko opremo). Ločimo med:
  - migracijo verzije (npr. SGML > HTML > XML),
  - migracijo formata (sprememba logičnega formata),
  - migracija dostopnosti,
  - reinženiring softvera (nove platforme in operacijski sistemi);
- **emulacija** (dodatki v hardveru/softveru, ki omogočajo, da se določena tehnologija obnaša kot druga - npr. novi operacijski sistem - dokument obdrži originalno logično kodo;
- **nedigitalne metode** (tiskanje, mikrofilmiranje ...);
- **kombinacija strategij** (tiskanje in shranjevanje v digitalni obliki);
- **druge metode** (npr. abstrakcija podatkov - UVC - *Universal Virtual Computer*) (Lorie, R.A., 2001);
- **povrnitev podatkov** (*data archeology*).

Če je prioriteta ohranjanje fizičnega nivoja, bi morali poskrbeti, da ohranimo tehnološko okolje, v katerem je digitalni vir bil berljiv, ali razviti emulacijo, ki nam bo omogočila dostop do digitalnega vira brez spremembe njegovih lastnosti. Vendar ne smemo pozabiti, da je s časom vse težje vzdrževati staro informacijsko tehnologijo, ker so nadomestni deli in znanje, ki spremlja to tehnologijo, vse manj dosegljivi.

Če je pa za nas pomembnejši konceptualni nivo, potem se bomo odločili za migracijo digitalnega vira v višjo različico ne glede na logični nivo.

Pri zahtevnejših virih pa bomo morali najti pravi način za ohranjanje večine specifičnih lastnosti.

Zanimive so izkušnje nacionalne knjižnice Nove Zelandije, kjer ohranjajo originalni digitalni dokument, ne glede na to, ali je še berljiv, in z njim še različico, ki je berljiva in je prilagojena novejši tehnologiji.

### 3.6 USKLAJENOST ARHIVSKEGA SISTEMA Z REFERENČNIM MODELOM OAIS

OAIS je bil leta 2003 sprejet za standard (ISO 14721:2003) pri izgradnji vseh digitalnih arhivov. Omenjeni standard definira OAIS kot "arhiv, ki je sestavljen iz organizacije ljudi in sistemov in ki je sprejel odgovornost, da trajno ohranja informacije in omogoča določeni skupnosti dostop do le-teh" (CCSDS, 2002, str. 1). OAIS natančno opisuje funkcionalni in informacijski model, ki predstavljata ogrodje za izgradnjo digitalnih arhivov, in se ne spušča v specifikacijo tehnologije, praktičnih standardov in metapodatkov. To je prepuščeno posamičnemu arhivu glede na vsebino gradiva in funkcijo, ki jo ima v svojem okolju.<sup>9</sup>

---

<sup>9</sup> Več o Teoretičnem modelu OAIS v Kavčič-Čolić, A., 2004.



### 3.7 IZBIRA PRIMERNIH METAPODATKOV

Na osnovi referenčnega sistema OAIS potrebujemo različne tipe metapodatkov za spremljanje digitalnih virov v arhivskem sistemu. Splošno znana delitev po funkciji loči metapodatke na:

- **Deskriptivne metapodatke** - poznamo jih tako v bibliotekarstvu kot v arhivistiki. Pomagajo nam pri opisu in identifikaciji vira v repozitoriju. Danes je že nekaj uveljavljenih struktur formatov, ki omogočajo opis virov, kot so npr. formati tipa Marc, Dublin Core ali EAD;
- **Strukturne metapodatke** (METS, MODS ...) - posredujejo informacije strukturi digitalnih virov. To je predvsem pomembno pri kompleksnih oblikah, na primer monografijah ali zbornikih oz. pri multimedijskih digitalnih virih;
- **Administrativne metapodatke** - uporabljajo se za upravljanje in administriranje digitalnih virov. Posredujejo informacije o nabavi, lokaciji, različici ipd.;
- **Tehnične metapodatke** - posredujejo informacije o privilegijih dostopa, formatih o delovanju sistema itd.;
- **Metapodatke za trajno ohranjanje**. Nobeden od omenjenih metapodatkov ne posreduje informacij, ki so pomembne za trajno ohranjanje digitalnih virov. To nalogo imajo metapodatki za trajno ohranjanje (*long-term preservation metadata*) vsebujejo informacije o formatu oz. o strukturi bitnega toka (*bit stream*) digitalnega vira, značilnosti aplikacije, s katero lahko odpremo digitalni dokument, informacije o njegovih lastnostih in funkcionalnostih, spremembe izvedene na njem itd.

V času od leta 2002-2004 je Narodna in univerzitetna knjižnica v sodelovanju z Institutom "Jožef Stefan" sodelovala v projektu "Metodologija zbiranja in arhiviranja slovenskih elektronskih publikacij na medmrežju" (ASEP). Projekt je sofinanciralo Ministrstvo za šolstvo, znanost in šport, Ministrstvo za kulturo in Ministrstvo za informacijsko družbo v okviru Ciljnega raziskovalnega programa "Konkurenčnost Slovenije 2001-2006". V okviru projekta smo izdelali izčrpno metodologijo arhiviranja digitalnih publikacij na medmrežju. Ena od pomembnih nalog je bila tudi izdelava osnovnega nabora metapodatkov za opis digitalnega gradiva, ki naj bi se ohranjalo za prihodnost. Takrat smo na osnovi mednarodnih izkušnjah na projektih NEDLIB, CEDARS in ugotovitvah v okviru skupine RLG/OCLC pripravili nabor metapodatkov za trajno ohranjanje digitalnih virov (Kavčič-Čolić, 2004/1). Vendar se je situacija od takrat nekoliko spremenila in mnoge metapodatke za trajno ohranjanje digitalnih virov, predvsem tehnične narave, je možno avtomatsko vnašati v arhivski sistem.

RLG in OCLC je ustanovil ekspertno mrežo PREMIS (*PREservation Metadata Implementation Strategies*)<sup>10</sup>, ki ima za cilj določiti minimalni nabor metapodatkov za trajno ohranjanje digitalnih virov. Pomemben je tudi vpliv, ki ga ima v Evropi nabor metapodatkov novozelandske nacionalne knjižnice (*Metadata standards ... 2002*). Ta upošteva različne nivoje granularnosti digitalnih virov od nivoja zbirke do nivoja posamičnih datotek, iz katerih je vir sestavljen.

<sup>10</sup> Opis projekta PREMIS na URL: <http://www.oclc.org/research/projects/pmwg/> (pridobljeno 7. 2. 2005).

#### 4. ZAKLJUČEK

Področje trajnega ohranjanja digitalnih virov je v eksperimentalnem razvoju. Trenutno lahko z gotovostjo trdimo, da nobena rešitev ni idealna. Le čas nam bo pokazal, ali smo zgrešili cilj. Vendar, različni projekti in poskusi na nacionalni in mednarodni ravni kažejo, da je potrebno čimprej nekaj ukreniti in ne sedeti prekrizanih rok.

Zaskrbljujoče je, da so v zadnjih letih mnogi slovenski nacionalni programi podpirali dejavnosti s področja digitalizacije in gradnje digitalnih arhivov, vendar pred tem niso poskrbeli za ustvarjanje pogojev za ohranjanje tega gradiva. Kot smo videli v tem prispevku, mora nekdo prevzeti odgovornost in poskrbeti za ohranjanje za prihodnost.

Smernice pa so nastale na osnovi večletnih izkušenj institucij, ki so veliko prispevale k področju gradnje in razvoja digitalnih arhivov. Vsekakor nam dajejo izhodiščno točko pri odločanju in načrtovanju našega digitalnega arhiva, vendar je to lahko le en pristop, ki je nastal v specifičnih okoliščinah.

V Sloveniji bi bilo nujno potrebno ustanoviti mrežo strokovnjakov različnih profilov in iz različnih tipov organizacij, ki bi na nacionalni ravni skušali definirati standarde in priporočila za vse digitalne arhive. Podobne mreže se že ustanavljajo v tujini, na primer NESTOR (*Network of Expertise in Digital Preservation*) (Dobratz & Neuroth, 2004). Koordinatorji NESTOR-ja so večkrat izrazil željo in interes za sodelovanje z drugimi podobnimi nacionalnimi mrežami. V prihodnosti lahko pričakujemo vse več tovrstnega mednarodnega povezovanja. Sodelovanje s tujimi strokovnjaki in združenji nam lahko pomaga pri definiciji lastnega nacionalnega pristopa k gradnji digitalnih arhivov.

#### 5. LITERATURA

- *CCSDS (Consultative Committee for Space Data Systems)(2002): Reference Model of an Open Archival Information System (OAIS): Recommendation for space data system standards: CCSDS 650.0-B-1. Blue Book, January 2002. - Washington: CCSDS Secretariat. Pridobljeno 21. 9. 2004 s spletne strani [http://ssdoo.gsfc.nasa.gov/nost/isoas/ref\\_model.html](http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html).*
- *CENL & FEP (1999): Code of practice for the voluntary deposit of electronic publications. Pridobljeno 7. 2. 2004 s spletne strani <http://www.bl.uk/gabriel/fep/> (Partnerships & Projects).*
- *DINI-Certificate. Document and Publication Repositories (2003). Göttingen: Deutsche Initiative für Netzwerkinformation, October 2003. Pridobljeno 7. 2. 2004 s spletne strani <http://www.dini.de/documents/Zertifikat-en.pdf>.*
- *Dobratz, Susanne and Neuroth, Heike: nestor: Network of Expertise in Long-term STORage of Digital Resources - A Digital Preservation Initiative for Germany (2004), D-Lib Magazine, 10(4) April. Pridobljeno 7. 2. 2004 s spletne strani <http://www.dlib.org/dlib/april04/dobratz/04dobratz.html>.*
- *Garret, John and Waters, Donald: Preserving Digital Information. Report of the Task Force on Archiving of Digital Information (1996). Commission on Preservation and Access and RLG. Pridobljeno 7. 2. 2004 s spletne strani <http://www.rlg.org/legacy/ftpd/pub/archtf/final-report.pdf>.*
- *Guidelines for the preservation of digital heritage (2003) / prepared by the National Library of Australia. - Paris: UNESCO. Pridobljeno 21. 9. 2004 s spletne strani <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>.*
- *Kavčič-Čolić, Alenka (2004/1): "Metapodatki za trajno ohranjanje elektronskih virov". Knjižnica 48(4), 2004, str. 97-119.*
- *Kavčič-Čolić, Alenka (2004/2): "Teoretični model digitalnih arhivov". Knjižnica 48(4), 2004, str. 63-75.*
- *Lorie, Raymond A. (2001): "A Project on Preservation of Digital Data". RLG DigiNews, 5(3), 15 June. Pridobljeno 5. 5. 2004 s spletne strani <http://www.rlg.org/preserv/diginews/diginews5-3.html#feature2>.*

- *Lupovici, Catherine and Masanès, Julien (2000): Metadata for Long Term Preservation. - Bibliothèque Nationale de France. Den Haag: Koninklijke Bibliotheek. - (NEDLIB Report series ; 2). Pridobljeno 21. 9. 2004 s spletne strani <http://www.kb.nl/coop/nedlib/results/NEDLIBmetadata.pdf>.*
- *Metadata standards framework - preservation metadata (2002). Wellington: National Library of New Zealand. Pridobljeno 5.5.2004 s spletne strani [http://www.natlib.govt.nz/files/4initiatives\\_metaschema.pdf](http://www.natlib.govt.nz/files/4initiatives_metaschema.pdf).*
- *Research Libraries Group (RLG (2002): Trusted Digital Repositories: Attributes and responsibilities: an RLG-OCLC Report. Mountain View, CA: RLG. Pridobljeno 21. 9. 2004 s spletne strani <http://www.rlg.org/longterm/repositories.pdf>.*

## SUMMARY

### AN APPROACH TO THE DEVELOPMENT OF DIGITAL ARCHIVES

Due to the specific nature of digital publications, archivists are not only forced to think about the collection of these publications, but also about their preservation for the future. In connection with this many organizations, national libraries, archives as well as international societies and groups have been dealing with the issue of a long-term access to digital collections for the last few years. RLG (Research Library Group) and OCLC (Online Computer Library Center) reports defined the responsibilities and attributes of the so called trusted digital repositories (archives). According to these reports digital archives should fulfil certain conditions. Among others, their repositories should be in compliance with the Reference Model for an Open Archival Information System. Moreover, the organizations developing a digital archive should be financially and administratively trustworthy; they should provide an adequate information technology and professional staff for the development and maintenance of such collections; they should foresee possible long-term preservation strategies and changes in the archiving system; they should provide a high security for all systems; and they should be responsible for its overall performance. The development of digital archives should be planned in detail. A practical approach to its development was presented in the Guidelines for Preservation of Digital Heritage prepared for UNESCO by the National Library of Australia in 2003.

Based on international analyses and recommendations by RLG/OCLC and UNESCO Guidelines for Preservation of Digital Heritage an overview of the attributes of trusted digital archives is given and some aspects that should be taken into consideration in their development are presented. These aspects comprise: responsibility issues, selection criteria, cooperation with providers, copyright management, selection of long-term preservation strategies, metadata, etc.

*Alenka Kavčič-Čolić je začela študij španskega jezika in književnosti na Univerzi Complutense v Madridu, dokončala pa ga je na Univerzi v Beogradu leta 1983. Od leta 1988 deluje na področju knjižničarstva. Leta 1996 je magistrirala na področju informacijskih sistemov na Fakulteti za organizacijo in informatiko Univerze v Zagrebu. V Sloveniji so ji magistrsko diplomu priznali kot magisterij iz bibliotekarskih ved. V bibliotekarski stroki ima naziv bibliotekar specialist. Od leta 2001, ko je nastopila službo vodje Bibliotekarskega raziskovalnega centra NUK, se posveča arhiviranju elektronskih virov. Trenutno pripravlja disertacijo s področja arhiviranja spletnih strani na Oddelku za bibliotekarstvo Univerze v Zagrebu.*